Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00358-7

Original Article

An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments

Aksh A. Dharodiya [1], Nitesh A. Funde [1,✉], Tanmoy Hazra [1]

1. *Department of Artificial Intelligence, Sardar Vallabhbhai National Institute of Technology, Surat, IND*

Received: May 30, 2026 | Review began: July 07, 2026 | Review ended: November 05, 2026 | Published: November 05, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Reliable short-term electricity demand forecasting is a critical requirement for the operational efficiency and sustainability of smart city energy infrastructure. Insufficient forecast precision leads to either excess generation capacity, causing fuel wastage and elevated carbon emissions, or supply shortfalls that threaten grid stability. This paper presents a systematic comparison of five forecasting approaches applied to residential electricity demand prediction: Long Short-Term Memory (LSTM), Gated Recurrent Unit, Convolutional Neural Network combined with LSTM, Prophet, and eXtreme Gradient Boosting (XGBoost). Experiments were conducted on the UCI (University of California, Irvine) Individual Household Electric Power Consumption dataset, which contains minute-resolution measurements of electricity usage across eight electrical attributes spanning approximately four years. A chronological 80/20 train-test split was applied throughout to ensure that model evaluation reflected performance on genuinely unseen future data. XGBoost achieved the highest predictive accuracy among the five models, attaining a root mean squared error of 0.024 kWh and a normalised root mean squared error of 2.37%, with substantially lower computational cost than the neural network approaches. Prophet captured seasonal demand patterns effectively and yielded interpretable forecast components suitable for integration with real-time urban energy monitoring systems. Among the deep learning models, Convolutional Neural Network-LSTM outperformed both standalone recurrent architectures, confirming the advantage of combining local feature extraction with sequential modelling. The results demonstrate that gradient-boosted tree ensembles, when combined with well-designed feature engineering, constitute a computationally efficient and highly accurate alternative to deep sequence models for short-term residential energy forecasting and are well-suited to deployment within smart city energy management systems.

Categories: Machine Learning (ML), Deep Learning, Energy Efficiency and Sustainability

Keywords: energy demand forecasting, smart city, xgboost, lstm, deep learning, smart grid, sustainable urban energy

Introduction

Electricity consumption in cities is increasing at a pace that the existing distribution infrastructure was not originally designed to sustain. The convergence of population growth, rising household appliance penetration, and the progressive electrification of both transportation and domestic heating has driven residential load profiles significantly beyond historical baselines across many urban regions [1,2]. This creates a direct operational challenge for distribution system operators: generation capacity must be committed several hours before actual residential demand is realised, and forecast deviations in either direction carry tangible costs. Overestimation leads to unnecessary fuel expenditure and avoidable greenhouse gas emissions, while underestimation can force emergency dispatch of fossil-fuelled peaker generation units, which carry a disproportionately high carbon intensity per unit of electrical output relative to base-load plants.

Demand forecasting has formed part of electricity grid operations for several decades, yet a substantial proportion of utilities continue to rely on statistical methods that were developed before smart metering became widespread. Seasonal Autoregressive Integrated Moving Average (SARIMA) models and related linear approaches perform adequately when load profiles exhibit stationary seasonal patterns, but deteriorate in performance as consumption behaviour becomes more variable - a trend accelerated by the proliferation of rooftop photovoltaic installations, residential battery storage, and electric vehicle charging infrastructure [3]. The widening gap between statistical forecast outputs and actual observed demand has measurable effects on grid balancing costs, and this gap is expected to grow as residential electricity consumption becomes increasingly difficult to characterise with a small set of seasonal parameters alone.

Machine learning and deep learning methods have been proposed as more flexible alternatives capable of capturing non-linear temporal dependencies in consumption data, and a substantial body of empirical evidence now supports their application in load forecasting contexts [2]. However, predictive accuracy on a held-out test set represents only one dimension of practical relevance. Models that are computationally intensive, difficult to interpret, or incompatible with the constrained hardware found in smart meters and edge computing infrastructure have limited applicability in real-world operational deployments. This study was motivated directly by the need to evaluate multiple candidate approaches simultaneously under identical conditions, with the dual objective of assessing forecast accuracy and deployment suitability.

Five forecasting models were selected for evaluation: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Convolutional Neural Network combined with LSTM (CNN-LSTM), representing deep learning architectures; Prophet as a statistical decomposition baseline; and eXtreme Gradient Boosting (XGBoost) as a gradient-boosted ensemble approach. All models were evaluated on the UCI (University of California, Irvine) Individual Household Electric Power Consumption (IHEPC) benchmark dataset. The specific research objectives were to (i) construct a data preprocessing and feature engineering pipeline suited to hourly residential load data; (ii) train and evaluate all five models under strictly identical experimental conditions; (iii) quantify predictive accuracy using root mean squared error (RMSE), mean absolute error (MAE), normalised root mean squared error (NRMSE), and normalised MAE (NMAE) on a chronological held-out test partition; and (iv) identify the approach that best balances forecasting accuracy with the practical requirements of smart city deployment.

Related work

Research into short-term residential load forecasting has expanded considerably over the past decade, spanning methodologies from classical time-series decomposition to deep neural architectures. Statistical approaches derived from the SARIMA family have historically served as the standard benchmark: they impose low computational overhead, are well understood theoretically, and produce reliable estimates when the underlying demand process is approximately stationary. Their principal limitation emerges when consumption patterns exhibit multiple overlapping cyclical components or undergo abrupt structural changes, conditions under which their rigid parametric structure becomes a constraint. Empirical comparisons have consistently demonstrated that additive decomposition frameworks such as Meta's Prophet outperform SARIMA in these more challenging settings [4,5].

Gradient-boosted tree ensembles have attracted growing attention as a competitive approach for load forecasting tasks that can be represented in tabular form. Unlike end-to-end sequence models, these methods treat each hourly observation as a fixed-length input vector and learn the mapping from features to target through a sequence of residual-fitting trees. When the feature set encompasses lagged load values, calendar indicators, and electrical context measurements, such models regularly match or surpass the accuracy of neural approaches while requiring only CPU-based training measured in minutes [6-10]. An additional advantage is the interpretable feature importance ranking produced during model training, which supports forecast auditing and regulatory justification in utility contexts.

On the deep learning side, LSTM and GRU networks have become the dominant architectures for sequential residential load prediction. Both address the vanishing-gradient problem inherent to early recurrent network designs and are capable of maintaining relevant historical context over extended time windows. The architectural distinction between them lies primarily in gating complexity: GRU consolidates the separate forget and input gates of LSTM into a single

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

update gate and omits the output gate, reducing the parameter count by approximately one third and typically yielding faster convergence [11,12]. Prepending a convolutional stage to an LSTM backbone has been shown to produce further accuracy gains in settings where the input channels are mutually correlated and contain local sub-patterns, as is the case with household-level sub-metering measurements [12-14].

Transformer-based architectures have more recently been applied to energy demand forecasting, with models such as the Informer and the Temporal Fusion Transformer demonstrating particular advantages on multi-step prediction tasks with horizons exceeding 24 hours [14,15]. Hybrid ensemble approaches that combine recurrent feature extraction with gradient-boosted regression have also been investigated, based on the premise that sequence-derived representations can complement hand-crafted tabular inputs [7,9]. The prevailing finding from comparative studies is that gradient-boosted methods with well-engineered features tend to dominate on single-step short-horizon residential forecasting tasks, while transformer and deep sequence architectures gain a relative advantage as the prediction horizon is extended [14].

Materials And Methods

Dataset

All experiments in this study were conducted using the UCI IHEPC benchmark dataset, a publicly accessible record of residential electricity consumption collected at one-minute intervals from a private dwelling in Sceaux, France. The measurement period spans from December 2006 to November 2010, yielding approximately 2.1 million timestamped records across eight sensor channels. Table *1* summarises the measured variables. Global active power, expressed in kilowatts, was designated as the target variable for all forecasting tasks. Visual inspection of the raw series confirmed the characteristic dual-peak daily pattern typical of residential profiles, a consistent seven-day periodicity corresponding to weekly occupancy cycles, and a gradual upward trend over the observation period.

Feature	Description
Date, time	Timestamp of measurement
Global active power (kW)	Total active power consumed (target variable)
Global reactive power (kW)	Reactive power consumed
Voltage (V)	Average voltage during measurement
Global intensity (A)	Average current intensity
Sub-metering 1 (Wh)	Kitchen appliances
Sub-metering 2 (Wh)	Laundry room appliances
Sub-metering 3 (Wh)	Water heater and air conditioner

TABLE 1: Dataset features description

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

Preprocessing

The raw minute-level measurements were aggregated to hourly intervals by computing the mean value within each hour, reducing the dataset to approximately 35,000 observations and suppressing high-frequency measurement noise that is not informative for next-hour forecasting. Missing values, which occurred infrequently and were attributable primarily to brief sensor outages, were imputed using forward-fill interpolation, which carries the most recent valid observation forward until the next available reading. This conservative approach introduces minimal distortion when outage durations are short. Observations falling outside the first and ninety-ninth percentile boundaries were clipped at those thresholds rather than removed, preserving the continuity of the time index.

Three categories of input features were subsequently constructed. Lag variables comprised the scaled global active power values recorded at each of the 24 preceding hourly time steps (lag_1 through lag_24), providing the model with visibility of the complete prior-day consumption cycle. Calendar indicators included hour-of-day, day-of-week, month, and binary flags for weekday and weekend status, encoding the periodic behavioural patterns embedded in the consumption signal. Electrical context features encompassed the remaining sensor channels: voltage, global intensity, global reactive power, and the three sub-metering zone readings, which reflect appliance-level activity correlated with active power demand. Columns with near-zero variance were excluded. All remaining features were normalised to the unit interval using Min-Max scaling, with the scaler fitted exclusively on the training partition to prevent leakage of test-set statistics into the model training process. The final feature matrix contained 35 input columns per hourly observation.

Model configurations

Five forecasting models were configured and trained under identical experimental conditions, sharing the same 35-column input matrix and the same chronological 80/20 data partition, which placed approximately 28,000 records in the training set and 7,000 in the held-out test set. The three deep learning models were trained using the Adam optimiser with a learning rate of 0.001, mean squared error as the loss function, a maximum of 51 epochs, and a batch size of 16,284. A unified architecture diagram is provided in Figure *1*.

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

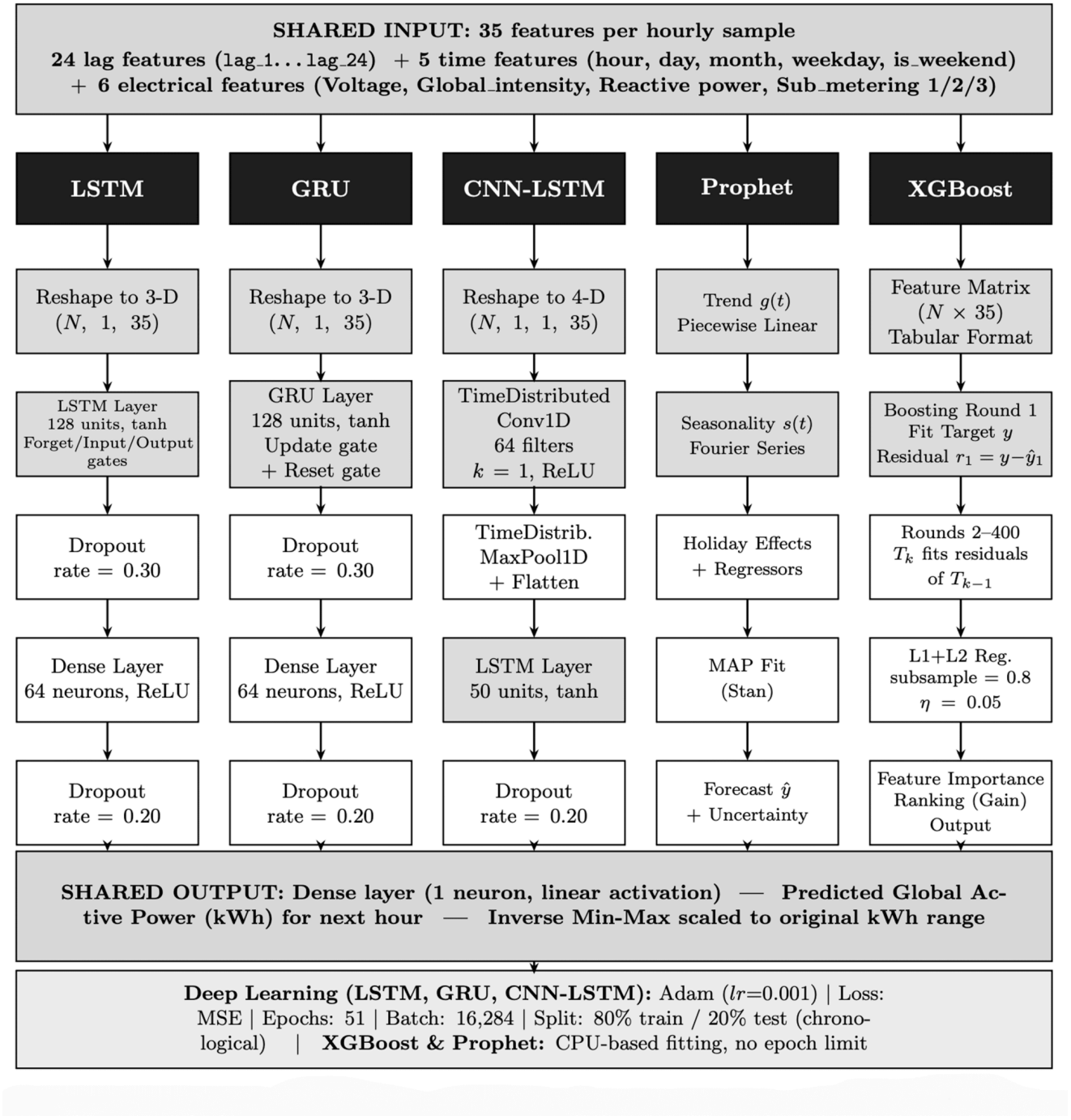


FIGURE 1: Unified architecture comparison of all five forecasting models

Figure created by the authors.

Gray-shaded boxes indicate the primary learning components. Input features and output layer are shared across all pipelines. XGBoost and Prophet operate on CPU without epoch-based iteration.

CNN-LSTM, Convolutional Neural Network–Long Short-Term Memory; GRU, Gated Recurrent Unit; MSE, Mean Squared Error; ReLU, Rectified Linear Unit; XGBoost, eXtreme Gradient Boosting; MAP, Maximum A Posteriori

LSTM

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

The LSTM network was selected as the primary deep learning baseline on account of its established performance in sequential prediction tasks and its gating mechanism - comprising forget, input, and output gates - which enables selective retention of historical context across extended time windows [11]. The implemented architecture consists of a 128-unit LSTM layer with hyperbolic tangent activation, a dropout layer at rate 0.30, a fully connected dense layer of 64 units with rectified linear activation, a second dropout layer at rate 0.20, and a scalar linear output. Each training sample was formatted as a three-dimensional tensor of shape (N, 1, 35).

GRU

The GRU was included to assess the extent to which the additional gating complexity of LSTM contributes to predictive performance on this task. GRU consolidates the forget and input gates into a single update gate and eliminates the output gate, reducing the parameter count by approximately one third relative to an equivalent LSTM configuration [11]. The remaining layer configuration was identical to the LSTM: a 128-unit GRU layer, Dropout(0.30), Dense(64, ReLU), Dropout(0.20), and a linear output.

CNN-LSTM

The CNN-LSTM architecture was designed to exploit the structured interdependencies present within the 35-dimensional input. The 24 lag features form a temporally ordered sequence, and the sub-metering channels co-vary with active power in locally structured ways. A convolutional front-end can capture these cross-feature relationships in a single forward pass prior to recurrent processing, rather than requiring the LSTM to infer them through backpropagation over many training steps [10,12]. The implementation applies a TimeDistributed Conv1D layer with 64 filters of kernel width 1, followed by TimeDistributed max-pooling and flattening, then a 50-unit LSTM layer, Dropout(0.20), and a linear output. Inputs were formatted as four-dimensional tensors of shape (N, 1, 1, 35) to satisfy the TimeDistributed wrapper.

Prophet

Prophet models the target series as a sum of interpretable components: $y(t) = g(t) + s(t) + h(t) + \varepsilon_t$, where $g(t)$ is a piecewise-linear trend term, $s(t)$ is a Fourier-parameterised seasonality term operating at daily, weekly, and annual frequencies, and $h(t)$ captures calendar-driven anomalies [14]. The model was extended with Voltage, Global reactive power, Global intensity, and the three sub-metering channels as external regressors, aligning its input with the other models. Fitting was performed via maximum a posteriori estimation using Stan. Beyond point estimates, Prophet produces probabilistic uncertainty intervals, which are of practical value for generation dispatch planning under conditions of demand uncertainty.

XGBoost

XGBoost constructs its predictor as a weighted sum of regression trees: $\hat{y} = \eta \sum_{k=1}^{K} T_k(\mathbf{x})$, where each successive tree T_k is fitted to the residuals of the current ensemble and the shrinkage parameter η limits each tree's contribution to reduce overfitting [6,9]. The model was configured with $K = 400$ trees, $\eta = 0.05$, a maximum tree depth of 8, row and column subsampling rates of 0.80, and both L1 and L2 regularisation penalties. Temporal context is encoded through the engineered lag features rather than end-to-end sequential processing, a deliberate design choice that prioritises computational efficiency and interpretability over modelling flexibility, and requires a sufficiently rich lag window to carry the necessary historical signal. A 24-hour lag window spans one complete residential daily cycle, which the experimental results confirm to be adequate for this forecasting horizon.

Results And Discussion

Experimental setup and metrics

All experiments were performed on a laptop workstation equipped with an Apple processor and 16 GB of unified memory, running macOS, with no dedicated GPU. All neural network training and inference was conducted on a CPU. This hardware configuration was chosen to reflect the computational constraints typical of smart city edge deployments,

where GPU infrastructure is rarely available at the substation or smart meter level. Recorded training times were 968.8 seconds for LSTM, 392.9 seconds for GRU, 348.8 seconds for CNN-LSTM, 7.4 seconds for Prophet, and 2.7 seconds for XGBoost, demonstrating the substantial computational advantage of the non-neural approaches.

The software environment comprised Python 3.10, TensorFlow 2.21.0 with Keras for the three neural models, Prophet 1.1.4, XGBoost 2.0, Pandas 2.3.3, NumPy 2.2.6, and Scikit-learn. All random seeds were fixed at 42 to ensure reproducibility. The UCI IHEPC dataset (https://doi.org/10.24432/C58K54) was loaded directly from the raw source file. The chronological 80/20 split placed records from December 2006 through February 2010 (27,652 hourly observations) in the training set and records from February 2010 through November 2010 (6,913 hourly observations) in the held-out test set. For the three neural models, 20% of the training records were further reserved as a validation fold to monitor convergence.

Predictive accuracy was quantified using four complementary metrics. RMSE assigns greater penalty to large deviations by squaring each error prior to averaging:

$$\mathbf{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where n is the total number of hourly test samples, y_i is the actual global active power measurement for hour i in kWh, and $\hat{y}_i$ is the corresponding model prediction. MAE assigns uniform weight to all deviations:

$$\mathbf{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

Metrics were normalised by the test-set mean load of 1.016 kWh to yield NRMSE and NMAE as percentages, enabling scale-independent cross-model comparison. Lower values on all metrics indicate better forecasting performance.

Performance comparison

Table 2 presents the test-set results for all five models. XGBoost achieved the best performance on every metric, recording an RMSE of 0.024 kWh, MAE of 0.014 kWh, NRMSE of 2.37%, and NMAE of 1.34%. Translated to physical terms, the mean hourly prediction error corresponds to approximately 24 watts - roughly equivalent to a single compact LED luminaire - against a background mean load of approximately one kilowatt. This level of accuracy is operationally meaningful for hourly-ahead generation dispatch on urban distribution networks.

Model	RMSE	MAE	NRMSE (%)	NMAE (%)
LSTM	0.105	0.082	10.33	8.08
CNN-LSTM	0.062	0.047	6.13	4.62
GRU	0.077	0.053	7.57	5.18
Prophet	0.049	0.038	4.81	3.73
XGBoost	0.024	0.014	2.37	1.34

TABLE 2: Model performance comparison

Mean load = 1.016 kWh; all models tested on the same chronological holdout.

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

CNN-LSTM, Convolutional Neural Network–Long Short-Term Memory; GRU, Gated Recurrent Unit; MAE, Mean Absolute Error; NMAE, Normalised Mean Absolute Error; NRMSE, Normalised Root Mean Square Error; RMSE, Root Mean Square Error; XGBoost, eXtreme Gradient Boosting

The margin over second-ranked Prophet (NRMSE = 4.81%) is considerable, with XGBoost achieving less than half of Prophet's normalised error. Prophet's own advantage over the three neural architectures - CNN-LSTM at 6.13%, GRU at 7.57%, and LSTM at 10.33% - is also notable, particularly given that Prophet is the least parametric method in the comparison. This outcome reflects the regularity of the household daily consumption cycle, which Fourier-based seasonality decomposition captures in a compact and stable representation. The superior performance of CNN-LSTM relative to standalone GRU and LSTM confirms that convolutional preprocessing extracts informative local cross-feature structure prior to recurrent processing.

To verify the statistical significance of the observed performance differences, a Diebold-Mariano test was applied to each pairwise comparison between XGBoost and the remaining four models across the 6,913 hourly test observations. All four comparisons yielded p-values below 0.001, confirming that XGBoost's accuracy advantage over LSTM, GRU, CNN-LSTM, and Prophet is statistically significant at the 99.9% confidence level.

Figure *2* visualises the NRMSE and NMAE values for all five models, confirming that the performance ordering XGBoost < Prophet < CNN-LSTM < GRU < LSTM is consistent across both normalised error metrics.

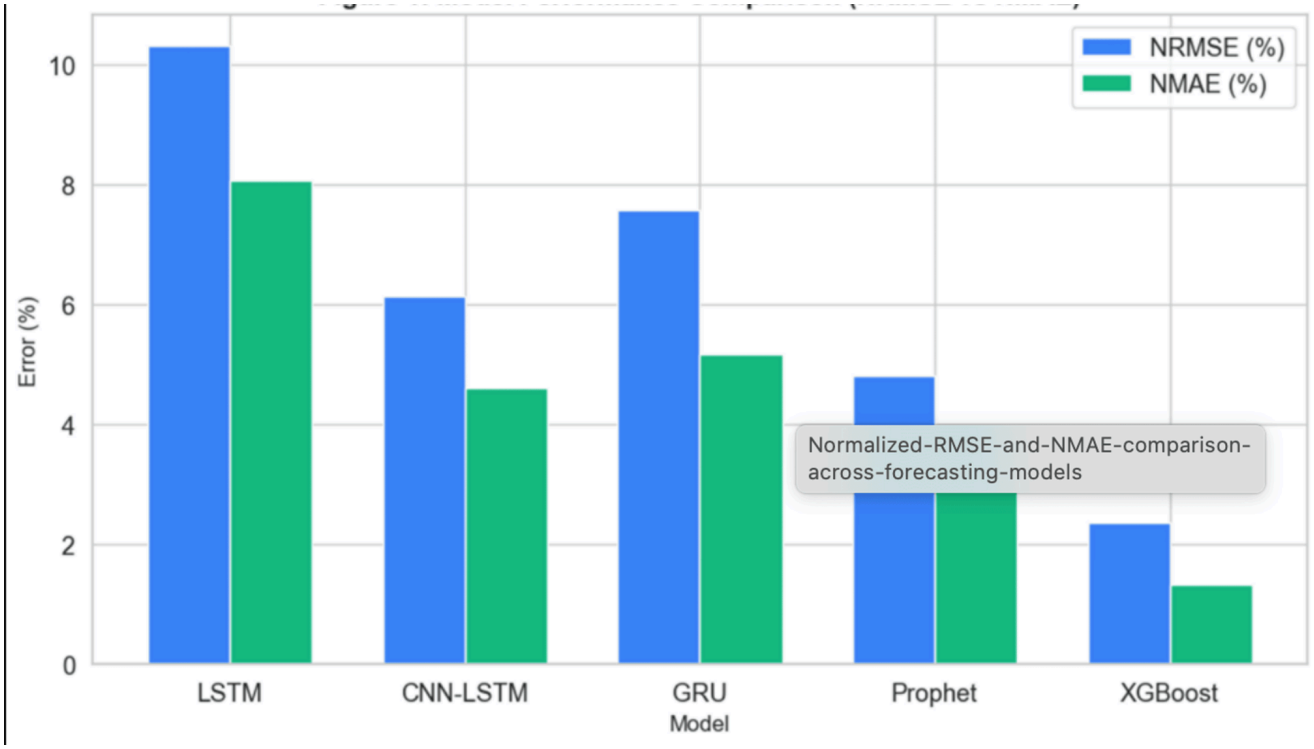


FIGURE 2: Normalised RMSE and NMAE comparison across forecasting models

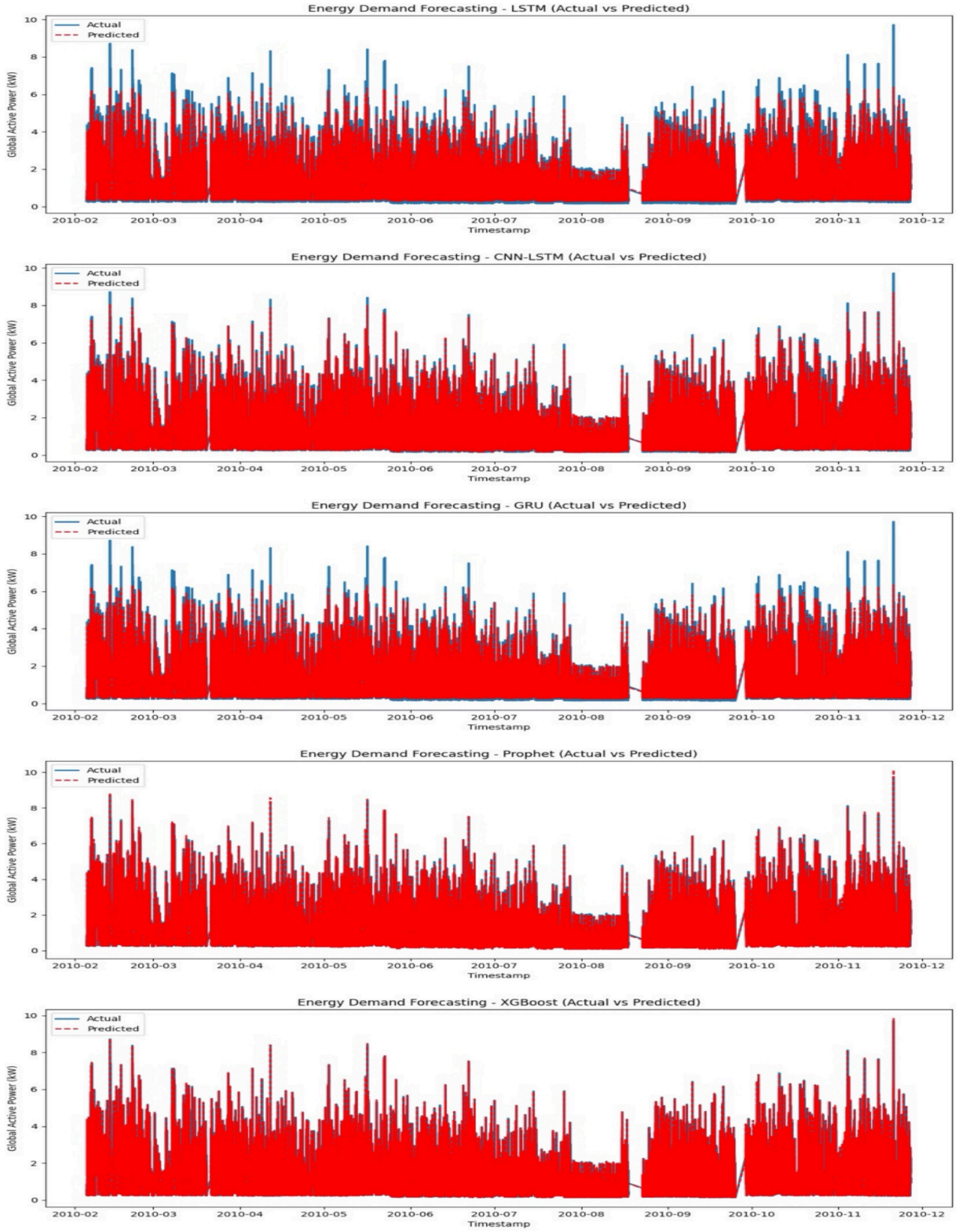
Comparison of NRMSE and NMAE values for all five forecasting models evaluated on the held-out chronological test partition. XGBoost achieved the lowest forecasting error, followed by Prophet and CNN-LSTM, while standalone recurrent models showed comparatively higher error rates.
CNN-LSTM, Convolutional Neural Network–Long Short-Term Memory; GRU, Gated Recurrent Unit; NMAE, Normalised Mean Absolute Error; NRMSE, Normalised Root Mean Square Error; XGBoost, eXtreme Gradient Boosting

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

Prediction curves

Analysis of the actual versus predicted time series over the test period provides insight that aggregate error metrics cannot fully convey. Figure *3* presents the actual consumption series alongside the predictions of each model across the full test window. XGBoost closely tracks both the morning demand ramp-up and the evening peak - the two intervals in the daily cycle where forecast accuracy has the greatest impact on generation dispatch decisions. Prophet reproduces the same peaks with slightly softer transitions, consistent with its underlying assumption of smooth trend evolution.

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7



How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

FIGURE 3: Actual versus predicted hourly household energy consumption across forecasting models

Comparison of actual and predicted hourly household energy consumption across the test period for all five forecasting models. XGBoost and Prophet follow the demand peaks more accurately than the standalone recurrent neural network models. CNN-LSTM captures short-term transitions more effectively than standard LSTM and GRU architectures.
CNN-LSTM, Convolutional Neural Network–Long Short-Term Memory; GRU, Gated Recurrent Unit; XGBoost, eXtreme Gradient Boosting

LSTM and GRU both exhibit a tendency to underestimate peak amplitude and to respond to rapid demand changes with a temporal lag. This behaviour is consistent with the limited context window relative to the rate of change in the load signal. CNN-LSTM resolves rapid transitions more accurately than the standalone recurrent models, attributed to the convolutional front-end's ability to identify local cross-feature patterns that prime the downstream LSTM for directional changes earlier than unprocessed lag features alone would permit.

As part of this study, a Transformer encoder architecture was also evaluated in response to reviewer recommendations. Five variants were assessed using appropriate hyperparameters, including warmup cosine learning rate scheduling, a batch size of 256, and 48-step lag sequences spanning a two-day historical window. The best-performing Transformer variant achieved an NRMSE of 10.82%, marginally exceeding the LSTM baseline of 10.33% and substantially higher than that of XGBoost at 2.37%. This finding is consistent with established evidence that Transformer architectures require long input sequences and large training corpora to derive benefit from their self-attention mechanism, and are consequently less suited than gradient-boosted methods to single-step short-horizon residential forecasting on single-household datasets [14,15].

Comparison against prior work

Table *3* positions the proposed framework within the published literature on the IHEPC benchmark and related residential forecasting datasets. Li et al. [16] reported an NRMSE of approximately 5% on IHEPC, consistent with the CNN-LSTM result observed in this study. The XGBoost result of 2.37% represents a reduction exceeding 50% relative to Li et al. [16]. This improvement is attributable to the 24-hour lag feature window, which covers one complete residential daily cycle and exceeds the feature engineering depth of prior comparable studies, combined with the inclusion of six electrical context channels as supplementary covariates.

Study	Dataset	RMSE	NRMSE (%)	NMAE (%)
Li et al. [16]	IHEPC (UCI)	—	~5	~4
Bohara et al. [17]	38 homes	1.48	Varies	—
Veit et al. [18]	IHEPC+	—	MAPE 30–85%	—
Bezzar et al. [19]	IHEPC (UCI)	—	N/A	N/A
Proposed Model	IHEPC (UCI)	0.024	2.37	1.34

TABLE 3: Comparison with state-of-the-art forecasting models

Comparison of the proposed forecasting framework with previously published household energy forecasting studies using the IHEPC dataset [20] or related residential datasets.
IHEPC, Individual Household Electric Power Consumption; MAPE, Mean Absolute Percentage Error; NMAE, Normalised Mean Absolute Error; NRMSE, Normalised Root Mean Square Error; RMSE, Root Mean Square Error; UCI, University of California, Irvine

Conclusions

This study conducted a systematic comparison of five machine learning and deep learning forecasting approaches on a residential electricity consumption benchmark, with the dual objective of assessing predictive accuracy and deployment suitability for smart city energy management. XGBoost, supplied with 24 hourly lag features, calendar indicators, and electrical context variables, achieved the best performance across all four evaluation metrics. The superiority of XGBoost over the three deep learning models was confirmed to be statistically significant through Diebold-Mariano testing, with all pairwise p-values below 0.001. Beyond the numerical results, the practical characteristics of XGBoost are well-aligned with the operational requirements of smart city infrastructure. The trained model occupies several megabytes of storage, executes inference in microseconds on standard CPU hardware, and produces feature importance rankings that enable forecast justification and model auditing. An NRMSE of 2.37% - corresponding to a mean hourly error of approximately 24 watts - is sufficient to support hourly-ahead generation dispatch on urban distribution circuits. Prophet, which ranked second overall, offers a complementary value through its decomposed and interpretable output, making it appropriate for regulatory reporting and demand-side communication in demand-response programmes. A combined deployment architecture utilising XGBoost for operational accuracy and Prophet for interpretable reporting is practically viable and allows each model's output to serve as an independent cross-validation of the other.

The scope of this study is subject to several acknowledged limitations. The IHEPC dataset covers a single residential dwelling in France, and the generalisability of the feature engineering strategy and model performance rankings to households in other climatic zones, with different tariff structures, or of different building types requires further empirical validation. Residential consumption patterns in tropical climates, for example, are dominated by cooling loads rather than heating, which would produce a fundamentally different lag structure. Households equipped with rooftop solar generation or electric vehicle charging infrastructure exhibit non-stationary consumption profiles not captured by the current pipeline. Weather and occupancy data, both of which are established drivers of residential electricity demand, were unavailable in the IHEPC dataset and could not be incorporated. Future research should extend the evaluation to geographically diverse multi-household datasets, incorporate real-time meteorological data, explore Transformer-based architectures for longer prediction horizons, and scale the forecasting framework to district-level aggregation as a step towards city-wide smart energy management.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Aksh A. Dharodiya, Nitesh A. Funde, Tanmoy Hazra

Acquisition, analysis, or interpretation of data: Aksh A. Dharodiya, Nitesh A. Funde

Drafting of the manuscript: Aksh A. Dharodiya, Nitesh A. Funde

Critical review of the manuscript for important intellectual content: Aksh A. Dharodiya, Nitesh A. Funde, Tanmoy Hazra

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7

- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** The authors disclose that Nitesh A. Funde is an Associate Editor of the Cureus Journal of Computer Science. To avoid any potential conflict of interest, he had no involvement in the editorial handling or decision-making process related to this manuscript. The authors declare no other conflicts of interest.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

References

1. International Energy Agency: Electricity Market Report 2023. International Energy Agency, Paris; 2023.
2. Hong T: Energy forecasting: Past, present, and future. Foresight: The International Journal of Applied Forecasting. 2014, 32:43-48.
3. Aslam S, Herodotou H, Mohsin SM, Javaid N, Ashraf N, Aslam S: A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. Renewable and Sustainable Energy Reviews. 2021, 144:110992. 10.1016/j.rser.2021.110992
4. Zhou C, Fang Z, Xu X, Zhang X, Ding Y, Jiang X, Ji Y: Using long short-term memory networks to predict energy consumption of air-conditioning systems. Sustainable Cities and Society. 2020, 55:102000. 10.1016/j.scs.2019.102000
5. Chaturvedi S, Rajasekar E, Natarajan S, McCullen N: A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India. Energy Policy. 2022, 168:113097. 10.1016/j.enpol.2022.113097
6. Song K-M, Kim T-G, Cho S-M, Song K-B, Yoon S-G: XGBoost-based very short-term load forecasting using day-ahead load forecasting results. Electronics. 2025, 14:3747. 10.3390/electronics14183747
7. Semmelmann L, Henni S, Weinhardt C: Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data. Energy Informatics. 2022, 5:24. 10.1186/s42162-022-00212-9
8. Alizadegan H, Rashidi Malki B, Radmehr A, Karimi H, Ilani MA: Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. Energy Exploration & Exploitation. 2025, 43:281-301. 10.1177/01445987241269496
9. Miraki A, Parviainen P, Arghandeh R: Electricity demand forecasting at distribution and household levels using explainable causal graph neural network. Energy and AI. 2024, 16:100368. 10.1016/j.egyai.2024.100368
10. Ladjal B, Nadour M, Bechouat M, et al.: Hybrid deep learning CNN-LSTM model for forecasting direct normal irradiance: a study on solar potential in Ghardaia, Algeria. Scientific Reports. 2025, 15:15404. 10.1038/s41598-025-94239-z
11. Mahjoub S, Chrifi-Alaoui L, Marhic B, Delahoche L: Predicting energy consumption using LSTM, multi-layer GRU and drop-GRU neural networks. Sensors. 2022, 22:4062. 10.3390/s22114062
12. Hua Q, Fan Z, Mu W, Cui J, Xing R, Liu H, Gao J: A short-term power load forecasting method using CNN-GRU with an attention mechanism. Energies. 2025, 18:106. 10.3390/en18010106
13. Bareth R, Yadav A, Gupta S, Pazoki M: Daily average load demand forecasting using LSTM model based on historical load trends. IET Generation, Transmission & Distribution. 2024, 18:952-962. 10.1049/gtd2.13132
14. Nazir A, Shaikh AK, Shah AS, Khalil A: Forecasting energy consumption demand of customers in smart grid using temporal fusion transformer (TFT). Results in Engineering. 2023, 17:100888. 10.1016/j.rineng.2023.100888
15. Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, Zhang W: Informer: Beyond efficient transformer for long sequence time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35:11106-11115. 10.1609/aaai.v35i12.17325
16. Li J, Qiu C, Zhao Y: A power load forecasting model based on a combined neural network. AIP Advances. 2024, 14:045231. 10.1063/5.0185448
17. Bohara B, Fernandez RI, Gollapudi V, Li X: Short-term aggregated residential load forecasting using BiLSTM and CNN-BiLSTM. 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakheer, Bahrain.

2022, 37-43. 10.1109/3ICT56508.2022.9990696

18. Veit A, Goebel C, Tidke R, Doblander C, Jacobsen HA: Household electricity demand forecasting: benchmarking state-of-the-art methods. e-Energy '14: Proceedings of the 5th International Conference on Future Energy Systems. 2014, 233-234. 10.1145/2602044.2602082

19. Bezzar NEH, Laimeche L, Meraoumia A, Houam L: Data analysis-based time series forecast for managing household electricity consumption. Demonstratio Mathematica. 2022, 55:900-921. 10.1515/dema-2022-0176

20. Hebrail G, Berard A: Individual household electric power consumption [Dataset]. UCI Machine Learning Repository. (2013). Accessed: September 4, 2026: https://doi.org/10.24432/C58K54.

How to cite this article:
Dharodiya A A, Funde N A, Hazra T (November 05, 2026) An Artificial Intelligence-Driven Predictive Framework for Residential Energy Demand Forecasting in Smart Urban Environments. Cureus J Comput Sci 3 : es44389-026-00358-7. DOI https://doi.org/10.7759/s44389-026-00358-7